

SEQUENCE LISTING

(1) GENERAL INFORMATION

(i) APPLICANT: SEED, BRIAN
HAAS, JURGEN

(ii) TITLE OF THE INVENTION: HIGH LEVEL EXPRESSION OF
PROTEINS

(iii) NUMBER OF SEQUENCES: 110

(iv) CORRESPONDENCE ADDRESS:

- (A) ADDRESSEE: Clark & Elbing LLP
- (B) STREET: 176 Federal Street
- (C) CITY: Boston
- (D) STATE: MA
- (E) COUNTRY: USA
- (F) ZIP: 02110

(v) COMPUTER READABLE FORM:

- (A) MEDIUM TYPE: Diskette
- (B) COMPUTER: IBM Compatible
- (C) OPERATING SYSTEM: DOS
- (D) SOFTWARE: FastSEQ for Windows Version 2.0

(vi) CURRENT APPLICATION DATA:

- (A) APPLICATION NUMBER: 08/717,294
- (B) FILING DATE: 20-SEP-1996
- (C) CLASSIFICATION:

(vii) PRIOR APPLICATION DATA:

- (A) APPLICATION NUMBER:
- (B) FILING DATE:

(viii) ATTORNEY/AGENT INFORMATION:

- (A) NAME: Elbing, Karen L
- (B) REGISTRATION NUMBER: 35,238
- (C) REFERENCE/DOCKET NUMBER: 00786/345001

(ix) TELECOMMUNICATION INFORMATION:

- (A) TELEPHONE: 617-428-0200
- (B) TELEFAX: 617-428-7045
- (C) TELEX:

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 24 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Other

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

CGCGGGCTAG CCACCGAGAA GCTG

24

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 195 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Other

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

ACCGAGAACG TGTGGGTGAC CGTGTACTAC GGCGTGCCCG TGTGGAAGAG AGGCCACCAC	60
CACCTGTTC TGCGCCAGCG ACGCCAAGGC GTACGACACC GAGGTGCACA ACGTGTGGC	120
CACCCAGGCG TGCGTCCCCA CCGACCCCAA CCCCCAGGAG GTGGAGCTCG TGAACGTGAC	180
CGAGAACTTC AACAT	195

(2) INFORMATION FOR SEQ ID NO:3:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 34 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Other

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

CCACCATGTT GTTCTTCCAC ATGTTGAAGT TCTC	34
---------------------------------------	----

(2) INFORMATION FOR SEQ ID NO:4:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 33 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Other

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

GACCGAGAAC TTCAACATGT GGAAGAACAA CAT	33
--------------------------------------	----

(2) INFORMATION FOR SEQ ID NO:5:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 192 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Other

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

TGGAAGAAC ACATGGTGG A GCAGATGCAT GAGGACATCA TCAGCCTGT GGACCAGAGC	60
CTGAAGCCCT GCGTGAAGCT GACCCCTGT GCGTGACCTG AACTGCACCG ACCTGAGGAA	120
CACCAACCAAC ACCAACACAG CACCGCCAAC AACAAACAGCA ACAGCGAGGG CACCATCAAG	180
GGCGGCGAGA TG	192

(2) INFORMATION FOR SEQ ID NO:6:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 33 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single

(D) TOPOLOGY: linear
(ii) MOLECULE TYPE: Other
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

GTTGAAGCTG CAGTTCTTCA TCTCGCCGCC CTT

33

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 31 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Other

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

GAAGAACTGC AGCTTCAACA TCACCACCAAG C

31

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 195 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Other

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

AACATCACCA CCAGCATCCG CGACAAGATG CAGAAGGAGT ACGCCCTGCT GTACAAGCTG
GATATCGTGA GCATCGACAA CGACAGCACC AGCTACCGCC TGATCTCCTG CAACACCAGC
GTGATCACCC AGGCCTGCCA CAAGATCAGC TTTCGAGCCA TCCCCATCCA CTACTGCGCC
CCCGCCGGCT TCGCC

60

120

180

195

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 30 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Other

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

GAACTTCTTG TCGGCGGCCGA AGCCGGCGGG

30

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 47 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Other

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

GCGCCCCCGC CGGCTTCGCC ATCCTGAAGT GCAACGACAA GAAGTTC

47

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 198 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Other

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

GCCGACAAGA AGTTCAGCGG CAAGGGCAGC TGCAAGAACG TGAGGCACCGT GCAGTGCACC	60
CACGGCATCC GGCCGGGTGGT GAGCACCCAG CTCCTGCTGA ACGGCAGCCT GGCCGAGGAG	120
GAGGTGGTGA TCCGCAGCGA GAACTTCACC GACAACGCCA AGACCATCAT CGTGCACCTG	180
AATGAGAGCG TGCAAGATC	198

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 34 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Other

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

AGTTGGGACG CGTGCAGTTG ATCTGCACGC TCTC

34

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 30 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Other

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

GAGAGCGTGCA GAGATCAACTG CACGCGTCCC

30

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 120 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Other

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

AACTGCACGC GTCCCAACTA CAACAAGCGC AAGCGCATCC ACATCGGCC CGGGCGCGCC	60
TTCTACACCA CCAAGAACAT CATCGGCACC ATCCTCCAGG CCCACTGCAA CATCTCTAGA	120

(2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 30 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Other

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

GTCGTTCCAC TTGGCTCTAG AGATGTTGCA

30

(2) INFORMATION FOR SEQ ID NO:16:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 29 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Other

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

GCAACATCTC TAGAGCCAAG TGGAACGAC

29

(2) INFORMATION FOR SEQ ID NO:17:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 131 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Other

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

GCCAAGTGGAA CGCACACCCT GCGCCAGATC GTGAGCAAGC TGAAGGAGCA GTTCAAGAAC
AAGACCATCG TTGTCACCAG AGCAGCGCG GCGACCCCGA GATCGTGATG CACAGCTTCA
ACTGCGGCGG C

60

120

131

(2) INFORMATION FOR SEQ ID NO:18:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 29 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Other

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

GCAGTAGAAG AATTGCCGC CGCAGTTGA

29

(2) INFORMATION FOR SEQ ID NO:19:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 29 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Other

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

TCAACTGCGG CGGCGAATTG TTCTACTGC

29

(2) INFORMATION FOR SEQ ID NO:20:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 195 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Other

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

GGCGAATTCT TCTACTGCAA CACCAGCCCC CTGTTCAACA GCACCTGGAA CGGCAACAAC
ACCTGGAACA ACACCACCGG CAGCAACAAC AATATTACCC TCCAGTGCAA GATCAAGCAG
ATCATCAACA TGTGGCAGGA GGTGGGCAAG GCCATGTACG CCCCCCCCAG CGAGGGCCAG
ATCCGGTGCA GCAGC

60
120
180
195

(2) INFORMATION FOR SEQ ID NO:21:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 40 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Other

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

GCAGACCGGT GATGTTGCTG CTGCACCGGA TCTGGCCCTC

40

(2) INFORMATION FOR SEQ ID NO:22:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 40 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Other

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

CGAGGGGCCAG ATCCGGTGCA GCAGAACAT CACCGGTCTG

40

(2) INFORMATION FOR SEQ ID NO:23:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 198 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Other

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

AACATCACCG GTCTGCTGCT GACCCCGCGAC GGCGGCAAGG ACACCGACAC CAACGACACC
GAAATCTTCC GCCCCGGCGG CGCGACATG CGCGACAACG GGAGATCTGA GCTGTACAAG
TACAAGGTGG TGACGATCGA GCCCCGGGC GTGGCCCCCA CCAAGGCCAA GCGCCGCGTG
GTGCAGCGCG AGAAGCGC

60
120
180
198

(2) INFORMATION FOR SEQ ID NO:24:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 38 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Other

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

CGCGGGCGGC CGCTTTAGCG CTTCTCGCGC TGCACCAC

38

(2) INFORMATION FOR SEQ ID NO:25:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 39 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Other

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

CGCGGGGGAT CCAAGCTTAC CATGATTCCA GTAATAAGT

39

(2) INFORMATION FOR SEQ ID NO:26:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 165 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Other

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

ATGAATCCAG TAATAAGTAT AACATTATTA TTAAGTGAT TACAAATGAG TAGAGGACAA
AGAGTAATAA GTTTAACAGC ATCTTAGTA AATCAAAATT TGAGATTAGA TTGTAGACAT
GAAAATAATA CAAATTTGCC AATACAACAT GAATTTTCAT TAACG

60

120

165

(2) INFORMATION FOR SEQ ID NO:27:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 36 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Other

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

CGCGGGGAAT TCACGCGTTA ATGAAAATTC ATGTTG

36

(2) INFORMATION FOR SEQ ID NO:28:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 30 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single

(D) TOPOLOGY: linear
(ii) MOLECULE TYPE: Other
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:28:

CGCGGATCCA CGCGTGAAAA AAAAAAAACAT

30

(2) INFORMATION FOR SEQ ID NO:29:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 150 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Other

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:

CGTGAAAAAA AAAAACATGT ATTAAGTGG A CATTAGGAG TACCAGAAC A TACATATAGA
AGTAGAGTAA ATTGTTTAG TGATAGATT C ATAAAAGTAT TAACATTAGC AAATTTACA
ACAAAAGATG AAGGAGATTA TATGTGTGAG

60
120
150

(2) INFORMATION FOR SEQ ID NO:30:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 30 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Other

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:30:

CGCGAATTCTG AGCTCACACA TATAATCTCC

30

(2) INFORMATION FOR SEQ ID NO:31:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 30 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Other

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:31:

CGCGGATCCG AGCTCAGAGT AAGTGGACAA

30

(2) INFORMATION FOR SEQ ID NO:32:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 170 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Other

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:32:

CTCAGAGTAA GTGGACAAAA TCCAACAAAGT AGTAATAAAA CAATAATGT AATAAGAGAT

60

AAATTAGTAA AATGTGAGGA ATAAGTTAT TAGTACAAAA TACAAGTTGG TTATTATTAT
TATTATAAG TTTAAGTTT TTACAAGCAA CAGATTAT AAGTTATGA

120
170

(2) INFORMATION FOR SEQ ID NO:33:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 36 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Other

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:33:

CGCGAATTG CGGCCGCTTC ATAAACTTAT AAAATC

36

(2) INFORMATION FOR SEQ ID NO:34:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1632 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:34:

CTCGAGATCC ATTGTGCTCT AAAGGAGATA CCCGGCCAGA CACCCTCACC	TGCGGTGCC	60
AGCTGCCAG GCTGAGGCAA GAGAAGGCCA GAAACCATGC CCATGGGGTC	TCTGCAACCG	120
CTGGCACCT TGTACCTGCT GGGGATGCTG GTCGCTCCG TGCTAGCCAC	CGAGAACGCTG	180
TGGGTGACCG TGTACTACGG CGTGCCCGTG TGGAAGGAGG CCACCAC	CCTGTTCTGC	240
GCCAGCGACG CCAAGGCGTA CGACACCGAG GTGCACAACG TGTGGGCCAC	CCAGGCGTGC	300
GTGCCACCG ACCCCAACCC CCAGGAGGTG GAGCTCGTGA ACGTGACCGA	GAACTTCAAC	360
ATGTGGAAGA ACAACATGGT GGAGCAGATG CATGAGGACA TCATCAGCCT	GTGGGACCAAG	420
AGCCTGAAGC CCTGCGTGA GCTGACCCCCC CTGTGCGTGA CCCTGAACTG	CACCGACCTG	480
AGGAACACCA CCAACACCAA CAACAGCACC GCCAACAAACA ACAGAACAG	CGAGGGCACC	540
ATCAAGGGCG GCGAGATGAA CAACTGCAGC TTCAACATCA CCACCGACAT	CCGCGACAAG	600
ATGCAGAAGG AGTACGCCCT GCTGTACAAG CTGGATATCG TGAGCATCGA	CAACGACAGC	660
ACCAGCTACC GCCTGATCTC CTGCAACACC AGCGTGATCA CCCAGGCCCTG	GCCCAAGATC	720
AGCTTCGAGC CCATCCCCAT CCACTACTGTC GCCCCCGCCG GCTTCGCCAT	CCTGAAGTGC	780
AACGACAAGA AGTTCAGCGG CAAGGGCAGC TGCAAGAACG TGAGCACCGT	GCAGTGCACC	840
CACGGCATCC GGGCGGTGGT GAGCACCCAG CTCTGCTGA ACGGCAGCCT	GGCCGAGGAG	900
GAGGTGGTGA TCCGCAGCGA GAACTTCACC GACAACGCCA AGACCATCAT	CGTCACCTG	960
AATGAGAGCG TGCAGATCAA CTGCACCGT CCCAACTACA ACAAGCGCAA GCGCATCCAC	1020	
ATCGGCCCCC GGGCGCCCTT CTACACCACC AAGAACATCA TCGGCACCAT	CCGCCAGGCC	1080
CACTGCAACA TCTCTAGAGC CAAGTGGAAC GACACCCCTG GCCAGATCGT	GAGCAAGCTG	1140
AAGGAGCAGT TCAAGAACAA GACCATCGTG TTCAACCCAGA GCAGCGGCCG	CGACCCCGAG	1200
ATCGTGTATGC ACAGCTTCAA CTGCGCGCGC GAATTCTTCT ACTGCAACAC	CAGCCCCCTG	1260
TTCAACAGCA CCTGGAACCG CAACACACCC TGGAACAAACA CCACCGGCAG	CAACAACAAT	1320
ATTACCTCC AGTGAAGAT CAAGCAGATC ATCAACATGT GGCAGGAGGT	GGGCAAGGCC	1380
ATGTACGCC CCCCATCGA GGGCCAGATC CGGTGCAGCA GCAACATCAC	CGGTCTGCTG	1440
CTGACCCCGCG AGGGCGGCAA GGACACCGAC ACCAACGACA CCGAAATCTT	CCGCCCGGC	1500
GGCGGGGACA TGCGCGACAA CTGGAGATCT GAGCTGTACA AGTACAAGGT	GGTGACGATC	1560
GAGCCCTGG GCGTGGCCCC CACCAAGGCC AAGCGCCGCG TGGTGCAGCG	CGAGAACGCG	1620
TAAAGCGGCC GC		1632

(2) INFORMATION FOR SEQ ID NO:35:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2481 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:35:

ACCGAGAACG	TGTGGGTGAC	CGTGTACTAC	GGCGTGCCCG	TGTGGAAGGA	GGCCACCACC	60
ACCCCTGTTCT	GCGCCAGCGA	CGCCAAGGCG	TACGACACCG	AGGTGCACAA	CGTGTGGGCC	120
ACCCAGGCGT	GCGTGCCCC	CGACCCCAAC	CCCCAGGAGG	TGGAGCTCGT	GAACGTGACC	180
GAGAACTTCA	ACATGTGGAA	GAACAACATG	CTGGAGCAGA	TGCATGAGGA	CATCATCAGC	240
CTGTGGGACC	AGAGCCTGAA	GCCCTGCGTG	AAGCTGACCC	CCCTGTGCGT	GACCCCTGAAC	300
TGCACCGACC	TGAGGAACAC	CACCAACACC	AACAACAGCA	CCGCCAACAA	CAACAGCAAC	360
AGCGAGGGCA	CCATCAAGGG	CGCGAGATG	AAGAACTGCA	GCTTCACAT	CACCACCAGC	420
ATCCGCGACA	AGATGCAGAA	GGAGTACGCC	CTGCTGTACA	AGCTGGATAT	CGTGAGCATT	480
CACAACGACA	GCACCAAGCTA	CCGCCTGATC	TCCTGCAACA	CCAGCGTGT	CACCCAGGCC	540
TGCCCCAAAGA	TCAGCTTCGA	GCCCATCCCC	ATCCACTACT	GCGCCCCCGC	CGGCTTCGCC	600
ATCCTGAAGT	GCAACGACAA	GAAGTTTCAGC	GGCAAGGGCA	GCTGCAAGAA	CGTGACCAACC	660
GTGCAGTGCA	CCCACCGGCAT	CCGGCCGGTG	GTGAGCACCC	AGCTCTGCT	GAACGGCAGC	720
CTGGCCGAGG	AGGAGGGTGT	GATCCGCAGC	GAGAACTTCA	CCGACAACGC	CAAGACCATC	780
ATCGTGACCC	TGAATGAGAG	CGTGCAGATC	AACTGCACGC	GTCCCAACTA	CAACAAGCGC	840
AAGCGCATCC	ACATCGGCC	CGGGCGCGCC	TTCTACACCA	CCAAGAACAT	CATCGGCACC	900
ATCCGCGCAGG	CCCACTGCAA	CATCTCTAGA	GCCAAGTGGA	ACGACAACCT	GCGCCAGATC	960
GTGAGCAAGC	TGAAGGAGCA	GTTCAAGAAC	AAGACCATCG	TGTTCAACCA	GAGCAGCGGC	1020
GGCAGCCCCG	AGATCGTGT	GCACAGCTTC	AACTGGGGCG	GCGAATTCTT	CTACTGCAAC	1080
ACCAGCCCCC	TGTTCAACAG	CACCTGGAAC	GGCAACAAACA	CCTGGAACAA	CACCACCGGC	1140
AGCAACAAACA	ATATTACCC	CCAGTGCAG	ATCAAGCAGA	TCATCACACAT	GTGGCAGGAG	1200
GTGGGCAAGG	CCATGTACGC	CCCCCCCCATC	GAGGGCCAGA	TCCGGTGCAG	CAGCAACATC	1260
ACCGGTCTGC	TGCTGACCCG	CGACGGCGGC	AAGGACACCG	ACACCAACGA	CACCGAAATC	1320
TTCCGCCCCG	GGGGCGCGA	CATCGCGAC	AACTGGAGAT	CTGAGCTGTA	CAAGTACAAG	1380
GTGGTGACGA	TCGAGCCCC	GGGCGTGGCC	CCCACCAAGG	CCAAGCGCCG	CGTGGTGCAG	1440
CGCGAGAACG	GGGGCGCCAT	CGGCGCCCTG	TTCCTGGGCT	TCCTGGGGGC	GGCGGGCAGC	1500
ACCATGGGGG	CCGCCAGCGT	GACCCGTACC	GTGCAGGCC	GCCTGCTCCT	GAGCGGCATC	1560
GTGCAGCAGC	AGAACAAACCT	CCTCCCGGCC	ATCGAGGCC	AGCAGCATAT	GCTCCAGCTC	1620
ACCGTGTGGG	GCATCAAGCA	GCTCCAGGCC	CGCGTGTG	CCGTTGGAGCG	CTACCTGAAAG	1680
GACCAGCAGC	TCCTGGGCTT	CTGGGGCTGC	TCCGGCAAGC	TGATCTGCAC	CACCACGGTA	1740
CCCTGGAACG	CCTCCTGGAG	CAACAAGAGC	CTGGACGACA	TCTGGAACAA	CATGACCTGG	1800
ATGCAGTGGG	AGCGCGAGAT	CGATAACTAC	ACCAGCCTGA	TCTACAGCCT	GCTGGAGAAG	1860
AGCCAGACCC	AGCAGGAGAA	GAACGAGCAG	GAGCTGCTGG	AGCTGGACAA	CTGGGCGAGC	1920
CTGTGGAACT	GGTTCGACAT	CACCAACTGG	CTGTGGTACA	TCAAAATCTT	CATCATGATT	1980
GTGGGCGGCC	TGGTGGGCCT	CCGCATCGTG	TTCGCGTGC	TGAGCATCGT	GAACCGCGTG	2040
CGCCAGGGCT	ACAGCCCCCT	GAGCCTCCAG	ACCCGGCCCC	CCGTGCGCGC	CGGGCCCGAC	2100
CGCCCCGAGG	GCATCGAGGA	GGAGGGCGGC	GAGCGCGACC	GCGACACCCAG	CGGCAGGCTC	2160
GTGCACGGCT	TCCTGGCGAT	CATCTGGGTC	GACCTCCGCA	GCCTGTTCCCT	GTTCAGCTAC	2220
CACCACCGCG	ACCTGCTGCT	GATCGCCGCC	CGCATCGTGG	AACTCCTAGG	CGCCCGCGGC	2280
TGGGAGGTGC	TGAAGTACTG	GTGGAACCTC	CTCCAGTATT	GGAGGCCAGGA	GCTGAAGTCC	2340
AGCGCCGTGA	GCCTGCTGAA	CGCCACCGCC	ATCGCCGTGG	CCGAGGGCAC	CGACCGCGTG	2400
ATCGAGGTGC	TCCAGAGGGC	CGGGAGGGCG	ATCCTGCACA	TCCCCACCCG	CATCCGCCAG	2460
GGGCTCGAGA	GGGCGCTGCT	G				2481

(2) INFORMATION FOR SEQ ID NO:36:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 486 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:36:

ATGAATCCAG	TAATAAGTAT	AACATTATTA	TTAAGTGTAT	TACAAATGAG	TAGAGGACAA	60
AGAGTAATAA	GTAAACAGC	ATGTTTAGTA	AATCAAATT	TGAGATTAGA	TTGTAGACAT	120
GAAAATAATA	CACCTTTGCC	AATACAACAT	GAATTTTCAT	TAACCGTGA	AAAAAAAAAA	180
CATGTATTAA	GTGGAACATT	AGGAGTACCA	GAACATACAT	ATAGAAGTAG	AGTAAATTG	240
TTTAGTGATA	GATTCTAAA	AGTATTAACA	TTAGCAAATT	TTACAACAAA	AGATGAAGGA	300

GATTATATGT	GTGAGCTAG	AGTAAGTGG	CAAATCCAA	CAAGTAGTAA	TAAAACAATA	360
AATGTAATAA	GAGATAAATT	AGTAAATGT	GGAGGAATAA	GTTCATTAGT	ACAAAATACA	420
AGTTGGTTAT	TATTATTATT	ATTAAGTTA	AGTTTTTAC	AAGAACAGA	TTTTATAAGT	480
TTATGA						486

(2) INFORMATION FOR SEQ ID NO:37:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 485 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:37:

ATGAACCCAG	TCATCAGCAT	CACTCTCCTG	CTTCAGTCT	TGCAGATGTC	CCGAGGACAG	60
AGGGTGATCA	GCTGACAGC	CTGCCTGGTG	AACAGAACCT	TCGACTGGAC	TGCCGTATG	120
AGAATAACAC	CAACTTGCCC	ATCCAGCATG	AGTTTCAGCCT	GACCCGAGAG	AAGAAGAACG	180
ACGTGCTGTC	AGGCACCCCTG	GGGGTCCCCG	AGCACACTTA	CCGCTCCCGC	GTCAACCTTT	240
TCAGTGACCG	CTTTATCAAG	GTCCTTACTC	TAGCCAACCT	GACCACCAAG	GATGAGGGCG	300
ACTACATGTG	TGAACTTCGA	GTCTCGGGCC	AGAATCCCCAC	AAGCTCCAAT	AAAACATATCA	360
ATGTGATCAG	AGACAAGCTG	GTCAAGTGTG	GTGGCATAAG	CCTGCTGGTT	AAAAACACTT	420
CCTGGCTGCT	GCTGCTCCTG	CTTCCCTCT	CCTTCCTCCA	AGCCACGGAC	TTCATTTCTC	480
TGTGA						485

(2) INFORMATION FOR SEQ ID NO:38:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 33 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Other

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:38:

CGCGGGGCTA GCGCAAAGAG TAATAAGTTT AAC

33

(2) INFORMATION FOR SEQ ID NO:39:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 28 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Other

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:39:

CGCGGATCCC TTGTATTTTG TACTAATA

28

(2) INFORMATION FOR SEQ ID NO:40:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 762 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Other

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:40:

GAATTCACGC	GTAAGCTTGC	CGCCACCATG	GTGAGCAAGG	GCGAGGGAGCT	GTTCACCGGG	60
GTGGTCCCCA	TCCCTGGTCA	GCTGGACGGC	GACGTGAACG	GCCACAAGTT	CAGCGTGTCC	120
GGCGAGGGCG	AGGGCGATGC	CACCTACGGC	AAGCTGACCC	TGAAGTTCAT	CTGCACCACC	180
GGCAAGCTGC	CCGTGCCCTG	GCCCCACCC	GTGACCACCT	TCAGCTACGG	CGTGCAGTGC	240
TTCAGCCGCT	ACCCCGACCA	CATGAAGCAG	CACGACTTCT	TCAAGTCCGC	CATGCCCGAA	300
GGCTACGTCC	AGGAGCGCAC	CATCTCTTC	AAGGACGACG	GCAACTACAA	GACCCGGGCC	360
GAGGTGAAGT	TCGAGGGCGA	CACCCCTGGTG	AACCGCATCG	AGCTGAAGGG	CATCGACTTC	420
AAGGAGGACG	GCAACATCC	GGGGCACAAG	CTGGAGTACA	ACTACAACAG	CCACAACGTC	480
TATATCATGG	CCGACAAAGCA	GAAGAACGGC	ATCAAGGTGA	ACTTCAAGAT	CCGCCACAAAC	540
ATCGAGGACG	GCAGCGTGC	GCTCGCCGAC	CACTACCAGC	AGAACACCCC	CATCGGCGAC	600
GGCCCGTGC	TGCTGCCCGA	CAACCACTAC	CTGAGCACCC	AGTCCGCCCT	GAGCAAAGAC	660
CCCAACGAGA	AGCGCGATCA	CATGGCCTG	CTGGAGTTCG	TGACCGCCGC	CGGGATCACT	720
CACGGCATGG	ACGAGCTGTA	CAAGTAAAGC	GGCCGCGGAT	CC		762

(2) INFORMATION FOR SEQ ID NO:41:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 4670 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:41:

AAGCTTAAAC	CATGCCCATG	GGGTCTCTGC	AACCGCTGGC	CACCTTGTAC	CTGCTGGGA	60
TGCTGGTCGC	TTCCGTGCTA	GCCGCCACCA	GAAGATACTA	CCTGGGTGCA	GTGGAACGT	120
CATGGGACTA	TATGCAAAGT	GATCTCGGTG	AGCTGCCTGT	GGACGCAAGA	TTTCCTCCTA	180
GAGTGCCAAA	ATCTTTTCCA	TTCAACACCT	CAGTCGTGTA	AAAAAGACT	CTGTTGTAG	240
AATTACAGGA	TCACCTTTTC	AACATCGCTA	AGCCAAGGCC	ACCCCTGGATG	GGTCTGCTAG	300
GTCCTACCAT	CCAGGCTGAG	GTTCATGATA	CAGTGGTCA	TACACTTAAG	AACATGGCTT	360
CCCATCCTGT	CAGTCTTCAT	GCTGTTGGTG	TATCCTACTG	GAAAGCTTCT	GAGGGAGCTG	420
AATATGATGA	TCAGACCAGT	CAAAGGGAGA	AAGAAGATGA	TAAAGTCTTC	CCTGGTGGAA	480
GCCATACATA	TGTCCTGGCAG	GTCCTGAAAG	AGAATGGTCC	AATGGCCTCT	GACCCACTGT	540
GCCTTACCTA	CTCATATCTT	TCTCATGTGG	ACCTGGTAAA	AGACTTGAAT	TCAGGCCTCA	600
TTGGAGCCCT	ACTAGTATGT	AGAGAAGGGA	GTCTGGCCAA	GGAAAAGACA	CAGACCTTGC	660
ACAAATTAT	ACTACTTTT	GCTGTATTG	ATGAAGGGAA	AAGTGGCAC	TCAGAAACAA	720
AGAACTCCTT	GATGCAGGAT	AGGGATGCTG	CATCTGCTCG	GGCCTGGCCT	AAAATGCACA	780
CAGTCATGG	TTATGTAAC	AGGTCTCTGC	CAGGTCTGAT	TGGATGCCAC	AGGAAATCAG	840
TCTATTGGCA	TGTGATTGGA	ATGGGCACCA	CTCCTGAAGT	GCACATCAATA	TTCCTCGAAG	900
GTCACACATT	TCTTGTGAGG	AACCATCGCC	AGGCCTCCCT	GGAAATCTCG	CCAATAACTT	960
TCCTTACTGC	TCAAACACTC	TTGATGGACC	TTGGACAGTT	TCTACTGTT	TGTCAATATCT	1020
CTTCCCACCA	ACATGATGGC	ATGGAAGCTT	ATGTCAAAGT	AGACAGCTGT	CCAGAGGAAC	1080
CCCAACTACG	AATGAAAAAT	AATGAAGAAG	CGGAAGACTA	TGATGATGAT	CTTACTGATT	1140
CTGAAATGGA	TGTGGTCAGG	TTTGATGATG	ACAACCTCTC	TTCCCTTATC	CAAATTGCT	1200
CAGTTGCCAA	GAAGCATCCT	AAAACCTGGG	TACATTACAT	TGCTGCTGAA	GAGGAGGACT	1260
GGGACTATGC	TCCCTTAGTC	CTCGCCCCCG	ATGACAGAAG	TTATAAAAGT	CAATATTGAA	1320
ACAATGGCCC	TCAGCGGATT	GGTAGGAAGT	ACAAAAAAAGT	CCGATTATG	GCATACACAG	1380
ATGAAACCTT	TAAGACTCGT	GAAGCTATT	AGCATGAATC	AGGAATCTT	GGACCTTAC	1440
TTTATGGGA	AGTTGGAGAC	ACACTGTTGA	TTATATTAA	GAATCAAGCA	AGCAGACCAT	1500
ATAACATCTA	CCCTCACGG	ATCACTGATG	TCCGTCC	GTATTCAAGG	AGATTACCAA	1560
AAGGTGAAA	ACATTTGAAG	GATTTCCAA	TTCTGCCAGG	AGAAATATT	AAATATAAAT	1620
GGACAGTGAC	TGTAGAAGAT	GGGCCAACTA	AATCAGATCC	TCGGTGCCTG	ACCCGCTATT	1680
ACTCTAGTT	CGTTAATATG	GAGAGAGATC	TAGCTTCAGG	ACTCATTGGC	CCTCTCCTCA	1740
TCTGCTACAA	AGAATCTGTA	GATCAAAGAG	GAAACCAGAT	AATGTCAGAC	AAGAGGAATG	1800
TCATCCTGTT	TTCTGTATT	GATGAGAAC	GAAGCTGGTA	CCTCACAGAG	AATATACAAC	1860
GCTTCTCCC	CAATCCAGCT	GGAGTGCAGC	TTGAGGATCC	AGAGTCCAA	GCCTCCAACA	1920
TCATGCACAG	CATCAATGGC	TATGTTTTG	ATAGTTTGCA	GTTGTCAGTT	TGTTTGCATG	1980
AGGTGGCATA	CTGGTACATT	CTAACGATTG	GAGCACAGAC	TGACTTCCTT	TCTGTCTTCT	2040
TCTCTGGATA	TACCTTCAA	CACAAATGG	TCTATGAAGA	CACACTCACC	CTATTCCCAT	2100
TCTCAGGAGA	AACTGTCTTC	ATGTGCATGG	AAAACCCAGG	TCTATGGATT	CTGGGGTGCC	2160
ACAACTCAGA	CTTCGGAAC	AGAGGCATGA	CCGCCTTACT	GAAGGTTCT	AGTTGTGACA	2220

AGAACACTGG	TGATTATTAC	GAGGACAGTT	ATGAAGATAT	TTCAGCATAAC	TTGCTGAGTA	2280
AAAACAATGC	CATTGAACCA	AGAACGTTCT	CCCAGAATTC	AAGACACCCCT	AGCACTAGGC	2340
AAAAGCAATT	TAATGCCACC	CCACCAGTCT	TGAAACGCCA	TCAACGGGAA	ATAACTCGTA	2400
CTACTCTTCA	GTCAGATCAA	GAGGAAATTG	ACTATGATGA	TACCATATCA	GTTGAAATGA	2460
AGAAGGAAGA	TTTGACATT	TATGATGAGG	ATGAAAATCA	GAGCCCCCGC	AGCTTCAAA	2520
AGAAAACACG	ACACTATTT	ATTGCTGCAG	TGGAGAGGCT	CTGGGATTAT	GGGATGAGTA	2580
GCTCCCAACA	TGTTCTAAGA	AACAGGGCTC	AGAGTGGCAG	TGTCCCTCAG	TTCAAGAAAG	2640
TTGTTTCCA	GGAATTACT	GATGGCTCCT	TTACTCAGCC	CTTATACCGT	GGAGAACTAA	2700
ATGAACATT	GGGACTCCTG	GGGCCATATA	TAAGAGCAGA	AGTTGAAGAT	AATATCATGG	2760
TAACCTTCAG	AAATCAGGCC	TCTCGCCCT	ATTCCCTCTA	TTCTAGCCTT	ATTTCTTATG	2820
AGGAAGATCA	GAGGCAAGGA	GCAGAACCTA	AAAAAAACTT	TGTCAAGCCT	AATGAAACCA	2880
AAACTTACTT	TTGGAAAGTG	CAACATCATA	TGGCACCCAC	TAAAGATGAG	TTTGAUTGCA	2940
AAGCCTGGC	TTATTCTCT	GATGTTGACC	TGGAAAAGA	TGTGCACTCA	GGCCTGATTG	3000
GACCCCTCT	GGCTGCCAC	ACTAACACAC	TGAACCCCTGC	TCATGGGAGA	CAAGTGACAG	3060
TACAGGAATT	TGCTCTGTT	TTCACCATCT	TTGATGAGAC	CAAAGCTGG	TACTTCACTG	3120
AAAATATGGA	AAGAAAATGC	AGGGCTCCCT	GCAATATCCA	GATGGAAGAT	CCCACCTTTA	3180
AAGAGAATTA	TCGCTTCCAT	GCAATCAATG	GCTACATAAT	GGATACACTA	CCTGGCTTAG	3240
TAATGGCTCA	GGATCAAAGG	ATTGATGGT	ATCTGCTCAG	CATGGGCAGC	AATGAAAACA	3300
TCCATTCTAT	TCATTCAGT	GGACATGTGT	TCACTGTACG	AAAAAAAGAG	GAGTATAAAA	3360
TGGCACTGTA	CAATCTCTAT	CCAGGTTGTT	TTGAGACAGT	GGAAATGTTA	CCATCCAAAG	3420
CTGGAATTG	GGGGTGGAA	TGCCTTATTG	GCGAGCATCT	ACATGCTGGG	ATGAGCACAC	3480
TTTTCTGGT	GTACAGCAAT	AAGTGTCAAG	CTCCCCCTGGG	AATGGCTTCT	GGACACATTA	3540
GAGATTTC	GATTACAGCT	TCAGGACAAT	ATGGACAGT	GGCCCCAAAG	CTGGCCAGAC	3600
TTCATTATTC	CGGATCAATC	AATGCCCTGA	GCACCAAGGA	GCCCTTTCT	TGGATCAAGG	3660
TGGATCTGTT	GGCACCAATG	ATTATTACAG	GCATCAAGAC	CCAGGGTGCC	CGTCAGAACT	3720
TCTCCAGCCT	CTACATCTCT	CAGTTATCA	TCATGTATAG	TCTTGATGGG	AAGAAGTGGC	3780
AGACTTATCG	AGGAAATTCC	ACTGGAAACCT	TAATGGTCTT	CTTGGCAAT	GTGGATTCTAT	3840
CTGGGATAAA	ACACAATATT	TTAACCCCTC	CAATTATTGC	TCGATACATC	CGTTTGCACC	3900
CAACTCATTA	TAGCATTTCG	AGCACTCTTC	GCATGGAGTT	GATGGGCTGT	GATTAAATA	3960
GTTGCAGCAT	GCCATTGGGA	ATGGAGAGTA	AAGCAATATC	AGATGCACAG	ATTACTGCTT	4020
CATCCTACTT	TACCAATATG	TTTGCCACCT	GGTCTCCTTC	AAAAGCTCGA	CTTCACCTCC	4080
AAGGGAGGAG	TAATGCCCTGG	AGACCTCAGG	TGAATAATCC	AAAAGAGTGG	CTGCAAGTGG	4140
ACTTCCAGAA	GACAATGAAA	GTCACAGGAG	TAACTACTCA	GGGAGTAAAA	TCTCTGCTTA	4200
CCAGCATGTA	TGTGAAGGAG	TTCCCTCATCT	CCAGCAGTCA	AGATGGCCAT	CAGTGGACTC	4260
TCTTTTTCA	GAATGGCAA	GTAAAGGTTT	TTCAGGGAAA	TCAAGACTCC	TTCACACCTG	4320
TGGTGAAC	TCTAGACCCA	CCGTTACTGA	CTCGCTACCT	TCGAATTAC	CCCCAGAGTT	4380
GGGTGCACCA	GATTGCCCTG	AGGATGGAGG	TTCTGGGCTG	CGAGGCACAG	GACCTCTACT	4440
GAGGGTGGCC	ACTGCAGCAC	CTGCCACTGC	CGTCACCTCT	CCCTCCCTAG	CTCCAGGGCA	4500
GTGTCCCTCC	CTGGCTTGCC	TTCTACCTTT	GTGCTAAATC	CTAGCAGACA	CTGCCTTGAA	4560
GCCTCCTGAA	TTAACTATCA	TCAGTCTGTC	ATTTCCTTGG	TGGGGGGCCA	GGAGGGTGCA	4620
TCCAATTAA	CTTAACCTCTT	ACCGTCGACC	TGCAGGCCCA	ACGGGGCCGC		4670

(2) INFORMATION FOR SEQ ID NO:42:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 4451 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:42:

AAGCTTAAAC	CATGCCCATG	GGGTCTCTGC	AACCGCTGGC	CACCTTGTAC	CTGCTGGGGA	60
TGCTGGTCGC	TTCCGTGCTA	GCCGCCACCC	GCCGCTACTA	CCTGGGCGCC	GTGGAGCTGT	120
CCTGGGACTA	CATGCAGAGC	GACCTGGCG	AGCTCCCCGT	GGACGCCCGC	TTCCCCCCCC	180
GCGTGCCTAA	GAGCTTCCCC	TTCAACACCA	GCGTGGTGT	CAAGAAAACC	CTGTTCGTGG	240
AGTTCACCGA	CCACCTGTTC	AACATTGCCA	AGCCGCGCCC	CCCCTGGATG	GGCCTGCTGG	300
GCCCCACCAT	CCAGGCCGAG	GTGTACGACA	CCGTGGTGT	CACCCCTGAAG	AACATGGCCA	360
GCCACCCCGT	CAGCCTGCAC	GCCGTGGCG	TGAGCTACTG	GAAGGCCAGC	GAGGGCGCCG	420
AGTACGACGA	CCAGACGTCC	CAGCGCGAGA	AGGAGGACGA	CAAGGTGTTC	CGGGGGGGGA	480
GCCACACCTA	CGTGTGGCAG	GTGCTTAAGG	AGAACGGCCC	TATGCCAGC	GACCCCTGT	540
GCCTGACCTA	CAGCTACCTG	AGCCACGTGG	ACCTGGTGA	GGATCTGAAC	AGCGGGCTGA	600
TCGGCGCCCT	GCTGGTGTGT	CGCGAGGGCA	GCCTGGCCAA	GGAGAAAACC	CAGACCTG	660

ACAAGTTCAT	CCTGCTGTTC	GCGGTGTTCG	ACGAGGGGAA	GAGCTGGCAC	AGCGAGACTA	720
AGAACAGCCT	GATGCAGGAC	CGCGACGCCG	CCAGCGCCCG	CGCCTGGCCC	AAGATGCACA	780
CCGTTAACGG	CTACGTGAAC	CGCAGCCTGC	CCGGCCTGAT	CGGCTGCCAC	CGCAAGAGCG	840
TGTACTGGCA	CGTCATCGGC	ATGGGCACCA	CCCCTGAGGT	GCACAGCATC	TTCCTGGAGG	900
GCCACACCTT	CCTGGTGCGC	AACCACCGCC	AGGCCAGCCT	GGAGATCAGC	CCCATCACCT	960
TCCTGACTGC	CCAGACCCCTG	CTGATGGACC	TAGGCCAGTT	CCTGCTGTTC	TGCCACATCA	1020
GCAGCCACCA	GCACGACGGC	ATGGAGGCTT	ACGTGAAGGT	GGACAGCTGC	CCCGAGGAGC	1080
CCCAGCTGCG	CATGAAGAAC	AACGAGGAGG	CCGAGGACTA	CGACGACGAC	CTGACCGACA	1140
GCGAGATGGA	TGTCGTACGC	TTCGACGACG	ACAACAGCCC	CAGCTTCATC	CAGATCCGCA	1200
GCGTGGCCAA	GAAGCACCCCT	AAGACCTGGG	TGCACTACAT	CGCCGCCGAG	GAGGAGGACT	1260
GGGACTACGC	CCCGCTAGTA	CTGGCCCCCG	ACGACCGCAG	CTACAAGAGC	CAGTACCTGA	1320
ACAACGGCCC	CCAGCGCATC	GGCCGCAAGT	ACAAGAAGGT	GCGCTTCATG	GCCTACACCG	1380
ACGAGACTTT	CAAGACCCGC	GAGGCCATCC	AGCACGAGTC	CGGCATCCTC	GGCCCCCTGC	1440
TGTACGGCGA	GGTGGGCGAC	ACCCCTGCTGA	TCATCTTCAA	GAACCAGGCC	AGCAGGCCCT	1500
ACAACATCTA	CCCCCACGGC	ATCACCGACG	TGCGCCCCCT	GTACAGCCGC	CGCCTGCCCA	1560
AGGGCGTGAA	GCACCTGAAG	GACTTCCCCA	TCCTGCCCCG	CGAGATCTTC	AAGTACAAGT	1620
GGACCGTGAC	CGTGGAGGAC	GGCCCCACCA	AGAGCGACCC	CCGCTGCCTG	ACCCGCTACT	1680
ACAGCAGCTT	CGTGAACATG	GAGCGCGACC	TGGCCTCCGG	ACTGATCGGC	CCCCTGCTGA	1740
TCTGCTACAA	GGAGAGCGTG	GACCAGCGCG	GCAACCAGAT	CATGAGCGAC	AAGCGCAACG	1800
TGATCCGTGTT	CAGCGTGTTC	GACGAGAAC	GCAGCTGGTA	TCTGACCGAG	ACATCCAGC	1860
GCTTCCGTGCC	CAACCCCGCT	GGCGTGCAGC	TGGAAGATCC	CGAGTTCCAG	GCCAGCAACA	1920
TCATGCACAG	CATCAACGGC	TACGTGTTCG	ACAGCCTGCA	GCTGAGCGTG	TGCCTGCATG	1980
AGGTGGGCTA	CTGGTACATC	CTGAGCATCG	GGC CCCAGAC	CGACTTCCTG	AGCGTGTCT	2040
TCTCCGGGTA	TACCTTCAAG	CACAAGATGG	TGTACGAGGA	CACCCCTGACC	CTGTTCCCC	2100
TCTCCGGCGA	GACTGTGTTC	ATGTCTATGG	AGAACCCCCGG	CCTGTGGATT	CTGGGCTGCC	2160
ACAACAGCGA	CTTCCGCAAC	CGCGGATGTA	CTGCCCTGCT	GAAAGTCTCC	AGCTGCGACA	2220
AGAACACCGG	CGACTACTAC	GAGGACAGCT	ACGAGGACAT	CTCCGCCTAC	CTGCTGTCCA	2280
AGAACAAACGC	CATCGAGCCC	CGCTCCTTCT	CCAAAACCTC	CCGCCACCCC	AGCACCGTC	2340
AGAAGCAGTT	CAACGCCAAC	CCCCCCGTGC	TGAAGCGCCA	CCAGCGCGAG	ATCACCCGCA	2400
CCACCCGTCA	AAGCGACCAG	GAGGAGATCG	ACTACGACGA	CACCATCAGC	GTGGAGATGA	2460
AGAAGGAGGA	CTTCGACATC	TACGACGAGG	ACGAGAACCA	GAGCCCCCGC	TCCTTCCAAA	2520
AGAAAACCCG	CCACTACTTC	ATCGCCGCCG	TGGAGCGCCT	GTGGGACTAC	GGCATGAGCA	2580
GCAGCCCCCA	CGTCCTGCGC	AACCGCGCCC	AGAGCGGCAG	CGTGCCTCCAG	TTCAAGAAGG	2640
TGGTGTTCGA	GGAGTTCAC	GACGGCAGCT	TCACCCAGCC	CCTGTACCGC	GGCGAGCTGA	2700
ACGAGCACCT	GGGCCTGCTC	GGCCCTTACA	TCCGCGCCGA	GGTGGAGGAC	AACATCATGG	2760
TGACCTCCG	CAACCAAGCC	TCCCAGCCCT	ACTCCTTCTA	CTCCTCCCTG	ATCAGCTACG	2820
AGGAGGACCA	GCGCCAGGGC	GCCGAGCCCC	GCAAGAACCT	CGTGAAGGCC	AACGAGACTA	2880
AGACCTACTT	CTGGAAGGTG	CAGCACCACA	TGGCCCCCAC	CAAGGACGAG	TTCGACTGCA	2940
AGGCCTGGC	CTACTTCAGC	GACGTGGACC	TGGAGAACCA	CGTGCACAGC	GGCCTGATCG	3000
GCCCCCTGCT	GGTGTGCCAC	ACCAACACCC	TGAACCCCCC	CCACGGGAGG	CAGGTGACTG	3060
TGCAGGAATT	TGCCCTGTTC	TTCACCATCT	TCGACGAGAC	TAAGAGCTGG	TACTTCACCG	3120
AGAACATGGA	GCGCAACTGC	CGCGCCCCCT	GCAACATCCA	GATGGAAGAT	CCCACCTTCA	3180
AGGAGAACTA	CCGCTTCCAC	GCCATCAACG	GCTACATCAT	GGACACCCCTG	CCCGGCCCTGG	3240
TGATGGGCCA	GGACCAAGCGC	ATCCGCTGGT	ACCTGCTGTC	TATGGGCAGC	AACGAGAACAA	3300
TCCACAGCAT	CCACTTCAGC	GGCCACGTTT	TCACCGTGC	CAAGAAGGAG	GAGTACAAGA	3360
TGGCCCTGTA	CAACCTGTAC	CCCAGCGTGT	TCGAGACTGT	GGAGATGTC	CCCAGCAAGG	3420
CCGGGATCTG	GCGCGTGGAG	TGCCTGATCG	GGAGGACACT	GCACGCCGGC	ATGAGCACCC	3480
TGTTCTGGT	GTACAGCAAC	AAAGTGGCAGA	CCCCCCTGGG	CATGCCAGC	GGCCACATCC	3540
GCGACTTCCA	GATCACCGCC	AGCGGCCAGT	ACGGCCAGTG	GGCTCCCAAG	CTGGCCCGCC	3600
TGCACTACAG	CGGCAGCATIC	AACGCCCTGGT	CGACCAAGGA	GCCCTTCTCC	TGGATCAAGG	3660
TGGACCTGCT	GGCCCCCATG	ATCATCCACG	GCATCAAGAC	CCAGGGCGCC	CGCCAGAAAGT	3720
TCAGCAGCCT	GTACATCAGC	CAGTCATCA	TCATGTACTC	TCTAGACGGC	AAGAAGTGGC	3780
AGACCTACCG	CGGCAACAGC	ACCGGCACCC	TGATGGTGT	CTTCGGCAAC	GTGGACAGCA	3840
GCGGCATCAA	GCACAACATC	TTCAACCCCC	CCATCATCGC	CCGCTACATC	CGCCTGCACC	3900
CCACCCACTA	CAGCATCCGC	AGCACCCCTGC	GCATGGAGCT	GATGGGCTGC	GACCTGAACA	3960
GCTGCAGCAT	GCCCCCTGGGC	ATGGAGAGCA	AGGCCATCAG	CGACGCCAG	ATCACCGCCT	4020
CCAGCTACTT	CACCAACATG	TTCGCCACCT	GGAGCCCCAG	CAAGGCCGC	CTGCACCTGC	4080
AGGGCCGAG	CAACGCCTGG	CGCCCCCAGG	TGAACAACCC	CAAGGAGTGG	CTGCAGGTGG	4140
ACTTCCAGAA	AACCATGAAG	GTGACTGGCG	TGACCACCCA	GGCGTCAAG	AGCCTGCTGA	4200
CCAGCATGTA	CGTGAAGGAG	TTCCTGATCA	GCAGCAGCCA	GGACGCCAC	CAGTGGACCC	4260
TGTTCTCCA	AAACGGCAAG	GTGAAGGTG	TCCAGGGCAA	CCAGGACAGC	TTCACACCGG	4320
TCGTGAACAG	CCTGGACCCC	CCCCTGCTGA	CCCGCTACCT	GCGCATCCAC	CCCCAGAGCT	4380
GGGTGCACCA	GATCGCCCTG	CGCATGGAGG	TGCTGGGCTG	CGAGGCCAG	GACCTGTACT	4440
GAAGCGCCG	C					4451

(2) INFORMATION FOR SEQ ID NO:43:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 18 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Other

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:43:

GGGGATCCTC ACGTCTCA

18

(2) INFORMATION FOR SEQ ID NO:44:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 33 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Other

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:44:

CTGCTTCTGA CGCGTGCTGG GGTGGCGGGA GTT

33

(2) INFORMATION FOR SEQ ID NO:45:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 21 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Other

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:45:

CTGCTGAAAG TCTCCAGCTG C

21

(2) INFORMATION FOR SEQ ID NO:46:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 33 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Other

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:46:

GGCAGGTGCT TAAGGAGAAC GGCCCTATGG CCA

33

(2) INFORMATION FOR SEQ ID NO:47:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 39 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Other